

Claims

What is claimed is:

1. A method, comprising:
 - receiving an unlabeled audio clip;
 - processing the unlabeled audio clip to extract an audio fingerprint;
 - determining a stored audio fingerprint that matches the extracted audio fingerprint; and
 - determining a labeled audio clip based on the stored audio fingerprint.
2. The method of claim 1, further comprising:
 - determining information about the labeled audio clip; and
 - providing the information to a user.
3. The method of claim 2, wherein the unlabeled audio clip is a song.
4. The method of claim 1, wherein processing the unlabeled audio clip to extract an audio fingerprint comprises:
 - receiving an audio signal representing the unlabeled audio clip;
 - down-sampling the received audio signal into a mono audio stream;
 - processing the down-sampled audio signal by generating frequency domain coefficients to produce one or more audio samples;
 - performing feature extraction of the one or more audio samples to produce a compact data representation; and
 - packing the compact data representation into one or more sub-fingerprints.

5. The method of claim 4, wherein processing the down-sampled audio signal by generating frequency domain coefficients to produce one or more audio samples comprises:

segmenting the down-sampled audio signal into one or more frames; and
performing inverse discrete cosine transform on the one or more frames.

6. The method of claim 5, wherein performing inverse discrete cosine transform on the one or more frames captures properties of the down-sampled audio signal.

7. The method of claim 4, wherein the received audio signal is uncompressed.

8. The method of claim 4, further comprising combining the one or more sub-fingerprints to create a fingerprint block.

9. The method of claim 4, wherein the received audio signal has a sample rate of 44.1 kHz and wherein down-sampling the received audio signal into a mono audio stream comprises down-sampling the received audio signal into a mono audio stream with a sampling rate of 5 kHz.

10. The method of claim 4, wherein the received audio signal has a sample rate of 48 kHz and where down-sampling the received audio signal into a mono audio stream comprises down-sampling the received audio signal into a mono audio stream with a sampling rate of 5 kHz.

11. The method of claim 4, wherein the sub-fingerprint is 32 bits.
12. A system, comprising:
 - an audio fingerprint generator; and
 - a database,
 - wherein the audio fingerprint generator receives an unlabeled audio clip and wherein the audio fingerprint generator processes the unlabeled audio clip to extract an audio fingerprint,
 - wherein the database determines a stored audio fingerprint that matches the extracted audio fingerprint and wherein the database determines a labeled audio clip based on the stored audio fingerprint.
13. The system of claim 12, wherein the database determines information about the labeled audio clip and wherein the database provides the information to a user.
14. The system of claim 13, wherein the unlabeled audio clip is a song.
15. The system of claim 12, wherein the audio fingerprint generator processes the unlabeled audio clip to extract an audio fingerprint by receiving an audio signal representing the unlabeled audio clip, down-sampling the received audio signal into a mono audio stream, processing the down-sampled audio signal by generating frequency domain coefficients to produce one or more audio samples, performing feature extraction of the one or more audio samples to produce a compact data representation and packing the compact data representation into one or more sub-fingerprints.

16. The system of claim 15, wherein the audio fingerprint generator processes the down-sampled audio signal by segmenting the down-sampled audio signal into one or more frames and performing inverse discrete cosine transform on the one or more frames.

17. The system of claim 16, wherein performing inverse discrete cosine transform on the one or more frames captures properties of the down-sampled audio signal.

18. The system of claim 15, wherein the received audio signal is uncompressed.

19. The system of claim 15, wherein the audio fingerprint generator combines the one or more sub-fingerprints to create a fingerprint block.

20. The system of claim 15, wherein the received audio signal has a sample rate of 44.1 kHz and wherein the audio fingerprint generator down-samples the received audio signal by down-sampling the received audio signal into a mono audio stream with a sampling rate of 5 kHz.

21. The system of claim 15, wherein the received audio signal has a sample rate of 48 kHz and wherein the audio fingerprint generator down-samples the received audio signal by down-sampling the received audio signal into a mono audio stream with a sampling rate of 5 kHz.

22. The system of claim 15, wherein the sub-fingerprint is 32 bits.

23. A machine-readable medium containing instructions which, when executed by a processing system, cause the processing system to perform a method, the method comprising:

receiving an unlabeled audio clip;

processing the unlabeled audio clip to extract an audio fingerprint;

determining a stored audio fingerprint that matches the extracted audio fingerprint; and

determining a labeled audio clip based on the stored audio fingerprint.

24. The machine-readable medium of claim 23, further comprising:

determining information about the labeled audio clip; and

providing the information to a user.

25. The machine-readable medium of claim 24, wherein the unlabeled audio clip is a song.

26. The machine-readable medium of claim 23, wherein processing the unlabeled audio clip to extract an audio fingerprint comprises:

receiving an audio signal representing the unlabeled audio clip;

down-sampling the received audio signal into a mono audio stream;

processing the down-sampled audio signal by generating frequency domain coefficients to produce one or more audio samples;

performing feature extraction of the one or more audio samples to produce a compact data representation; and

packing the compact data representation into one or more sub-fingerprints.

27. The machine-readable medium of claim 26, wherein processing the down-sampled audio signal by generating frequency domain coefficients to produce one or more audio samples comprises:

segmenting the down-sampled audio signal into one or more frames; and
performing inverse discrete cosine transform on the one or more frames.

28. The machine-readable medium of claim 27, wherein performing inverse discrete cosine transform on the one or more frames captures properties of the down-sampled audio signal.

29. The machine-readable medium of claim 26, wherein the received audio signal is uncompressed.

30. The machine-readable medium of claim 26, further comprising combining the one or more sub-fingerprints to create a fingerprint block.

31. The machine-readable medium of claim 26, wherein the received audio signal has a sample rate of 44.1 kHz and wherein down-sampling the received audio signal into a mono audio stream comprises down-sampling the received audio signal into a mono audio stream with a sampling rate of 5 kHz.

32. The machine-readable medium of claim 26, wherein the received audio signal has a sample rate of 48 kHz and where down-sampling the received audio signal into

a mono audio stream comprises down-sampling the received audio signal into a mono audio stream with a sampling rate of 5 kHz.

33. The machine-readable medium of claim 26, wherein the sub-fingerprint is 32 bits.